PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/29, 15/82, 15/11, 5/10, C12Q 1/68, C07K 14/415, 16/16, A01H 5/00

A1

(11) International Publication Number: WO 99/00501

(43) International Publication Date: 7 January 1999 (07.01.99)

(21) International Application Number: PCT/GB98/01861

(22) International Filing Date: 25 June 1998 (25.06.98)

(30) Priority Data:
9713430.8 25 June 1997 (25.06.97) GB

(71) Applicant (*for all designated States except US*): PLANT BIOSCIENCE LIMITED [GB/GB]; Norwich Research Park, Colney Lane, Norwich NR4 7UH (GB).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): GRAY, John, Clinton [GB/GB]; 47 Barons Way, Comberton, Cambridge CB3 7EQ (GB). WALKER, Amanda, Ruth [AU/GB]; Garden Cottage, Westwick, Oakington, Cambridge CB4 5AR (GB).

(74) Agents: KREMER, Simon, M. et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: MATERIALS AND METHODS RELATING TO A PLANT REGULATORY PROTEIN

(57) Abstract

Disclosed are nucleic acids encoding *TTG1* from *Arabidopsis*, which is believed to act upstream of an Arabidopsis R homologue in the pathways leading, *inter alia*, to trichome differentiation and anthocyanin synthesis. Also disclosed are variants and mutants of said sequence, vectors, host cells, transgenic plants, polypeptide expression products and other related materials, plus also methods of manipulating the phenotypic characteristics associated with *TTG1*.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Materials and Methods relating to a Plant Regulatory Protein

Field of the invention

This invention relates to materials and methods relating to a plant regulatory protein. More particularly, the invention relates to the cloning and expression of the *TTG1* gene of *Arabidopsis thaliana*, and homologues from other species, and manipulation and use of the gene in plants.

Background of the invention

The protein encoded by the *TTG1* (transparent testa, glabra) locus plays a central role in many pathways in *Arabidopsis thaliana*. Many of these pathways are confined to effects on the epidermal cell layer of different tissues. Mutations at the *TTG1* locus have a large range of pleiotropic effects (Koornneef (1981) Arabid Inform. Serv. 18 45-51). It is known that *ttg1* mutants have a glabrous phenotype with no leaf or stem hairs (trichomes) which are normally derived from the L1 layer of cells, the outer single layer of cells covering the meristem that differentiates into all epidermal cells of the leaf. No purple anthocyanin pigments are present in the seed coat leading to the yellow cotyledons being visible through the transparent testa. In the wild-type plant, anthocyanins are present in the hypocotyl of seedlings and in the stem and leaves of plants as they age, and are inducible by many forms of stress including by high light, poor nutrients or water stress. Mutants of the *ttg1* locus completely lack anthocyanins in the epidermis of leaves and stems (Koornneef (1981) Arabid Inform. Serv. 18 45-51). Tufts of mucilage are absent from *ttg1* mutant seeds and the seeds show no secretion of mucilage on imbibing, unlike wild-type plants (Koornneef (1981) Arabid Inform. Serv. 18 45-51). In wild-type plants, root hairs extend from root epidermal cells only

in files of cells that contact two underlying cortical cells, whereas in *ttg1* mutants extra root hairs occur in the atrichoblast cell files (Galway et al (1994) Dev. Biol. 166 740-754). Seeds of *ttg1* mutant plants do not require drying and cold treatments to germinate and exhibit an altered seed dormancy compared to ecotypes such as Landsberg *erecta* (Koornneef et al (1982) Theoret Appl Genet 61 385-393).

Several genetic loci involved in trichome, or leaf hair, differentiation and development have been described from Arabidopsis (Koornneef (1981) Arabid Inform. Serv. 18 45-51, Hülskamp et al. (1994) Cell 76 555-566). Three loci that play a role in the initiation of trichomes have been identified; these are *GL1* (glabra 1), *TTG1* (transparent testa glabra) and *TRY* (triptychon). Mutations at the *GL1* locus lead to hairless plants whereas the *TRY* locus affects the spacing of trichomes, which form clumps in *try* mutant plants (Hülskamp et al. (1994) Cell 76 555-566). GL1 is a MYB transcription factor (Marks and Feldmann (1989) Plant Cell 1 1043-1050, Oppenheimer et al (1991) Cell 67 483-493). Mutations at another glabra locus *GL2* have some features in common with *ttg1*, although *gl2* mutants have normal anthocyanin content and have rudimentary trichomes suggesting the fate of these cells has already been determined (Koornneef (1981) Arabid Inform. Serv. 18 45-51). They have an increased number of ectopic root hairs, although the atrichoblast cells resemble wild-type cells more than in the *ttg1* mutant indicating a role later in the development of the root epidermal cells (Masucci et al (1996) Development 122 1253-1260). The effects on the seed coat and mucilage are similar to that of the *ttg1* mutants (Koornneef (1981) Arabid Inform. Serv. 18 45-51). The *GL2* locus encodes another transcription factor, a homeodomain protein with a leucine zipper domain. By *in situ* hybridization, the *GL2* gene is expressed in developing trichomes (Rerie et al (1994) Genes Dev. 8

1388-1389) and in the atrichoblast cell files of the roots (Masucci et al (1996) Development 122 1253-1260). Transcript levels of *GL2* are greatly reduced in *ttg1* mutants, suggesting that the *TTG1* gene product is required for normal expression of *GL2* (Cristina et al (1996) Plant J 10 393-402).

In *ttg1* mutants the anthocyanin biosynthetic pathway is blocked at the dihydroflavonol-4-reductase (DFR) step because no *DFR* message is detected in these mutants (Shirley et al (1995) Plant J 8 659-671) whereas transcripts encoding chalcone synthase and chalcone isomerase are unaffected. This resembles the effect of *Delila* mutants in Antirrhinum (Martin et al (1991) Plant J 1 37-49). In maize the equivalent locus called the *R* gene affects the whole pathway from chalcone synthase (CS) onwards. Delila and R are both MYC-like transcription factors (Ludwig et al (1989) Proc Natl Acad Sci USA 86 7092-7096, Goodrich et al (1992) Cell 68 955-964). R has been shown to activate directly the transcription of several genes encoding anthocyanin biosynthetic enzymes in conjunction with a MYB transcription factor encoded by the *C1* gene in maize (Goff et al (1992) Genes Dev. 6 864-875). Complementation of a *ttg1* mutant by cauliflower mosaic virus 35S promoter-*R* constructs (Lloyd et al (1992) Science 258 1773-1775) was used to suggest that *TTG1* might encode an Arabidopsis R homologue. A further characterised transcription factor (Caprice [CPC] - see Wada et al 1997, Science 277, 1113-1116) may act in the opposite way to TTG1 in promoting root hair development, and possible reducing the trichome number.

The *TTG1* locus has been broadly mapped by Koornneef (Koornneef et al (1982) Theoret Appl Genet 61 385-393) to chromosome 5 between *ms1* and *ga3*. At position 31.5 the *ttg1* locus has been used as a phenotypic marker in many crosses.

Summary of the invention

The present inventors have identified the *TTG1* locus as a gene encoding a WD40 repeat protein by complementation of a *ttg1* mutant with genomic DNA and by sequencing the gene in several *ttg1* mutant alleles. The *TTG1* gene has now been cloned and sequenced and the inventors have demonstrated that it encodes a WD40 repeat protein with 7 repeat units. The 1.6 kb transcript is present in all major organs. The identification of the product of the *TTG1* locus as a WD40 repeat protein rules out the possibility that the protein acts as a transcription factor, unlike the products of the other genes, *GL1* and *GL2*, affecting trichome development. Additionally the TTG1 protein bears no resemblance to the maize *R* gene. The present inventors propose that the WD40 repeat protein is a component of a signal transduction pathway which regulates expression or action of downstream transcription factors, and in particular that *TTG1* acts upstream of an Arabidopsis R homologue in the pathways leading to trichome differentiation and anthocyanin synthesis.

The *TTG1* gene has a novel sequence. No *Arabidopsis* genes showing significant homology to *TTG1* were identified in public databases. A protein of unknown function showing 61% amino acid identity is encoded on chromosome 3 of Arabidopsis (orf10 in database accession number X98130), but transcripts of this gene do not cross hybridise with *TTG1* at high stringency on Northern blots. However, a region of the *TTG1* protein showed homology to an Expressed Sequence Tag (EST) of unknown function. The EST came from a cell suspension culture from Eco type Columbia (clone library AC16H).

Additionally the *TTG1* gene shows 87.5% similarity to the an11 gene from Petunia. This gene is discussed by Vetten et al (1997) Genes & Development 11: 1422-1434, Pub. Cold Spring Harbor Laboratory Press, which may have been published before the claimed priority date of the

present invention. Interestingly, the *an11* locus is described as controlling anthocyanin pigmentation and hence flower colour - but apparently does not exert the pleiotropic effects (e.g. trichomes, anthocyanin in other parts of the plant) of the *TTG1* gene which forms the basis of the present invention.

A genomic sequence encompassing Arabidopsis TTG1 has recently (after the priority date of the present invention) been put on a database under accession number AB010068.

Thus according to a first aspect of the present invention there is provided a nucleic acid molecule including a nucleotide sequence encoding a polypeptide with *TTG1* function. Those skilled in the art will appreciate that "*TTG1* function" may be used to refer to the ability to manipulate the phenotypic characteristics of plants as described below when its expression is altered like the *TTG1* gene of Arabidopsis thaliana.

Manipulation of the phenotypic characteristics of plants may be achieved by altering the expression of the *TTG1* gene (by increasing/decreasing expression or by mutation) or by interfering with the normal function of the TTG1 protein. Further, manipulation may be achieved by providing for the expression of a further homologous transcript which is able to interact with the expression of the *TTG1* gene in such a way as to either prevent translation of the transcript occurring or to boost the levels of transcripts being translated.

Examples of phenotypic characteristics that may be manipulated in accordance with the present invention are given below. Preferably at least 2, 3, 4, 5 or 6 or more of these characteristics are manipulated:

1. Trichomes (hairs) on aerial parts of plants: trichomes have a number of functions and the present invention provides a way to increase and decrease the number of trichomes on different organs to enhance their effectiveness. The increase or decrease in the number of

trichomas may be utilized in:

(i) insect protection: due to mechanical effects and to chemicals from glandular hairs. These could be increased, to increase insect protection, on leaves or on cotyledons which often do not have hairs. Protection of cotyledons from insect attack may allow faster seedling growth.

(ii) chemical production: glandular trichomes are involved in producing pheromones, antifeedants and other chemicals, including essential oils, which may be increased if the number of trichomes is increased.

(iii) protection in hot, dry climates: hairs form boundary layers for decreased water loss. Hairier plants may have an advantage in warmer climates. Hairs may also provide shade and protection for meristem in young seedlings, allowing faster seedling growth.

(iv) salt removal from leaves: the presence of salt glands would allow trichomes to sequester or secrete salt. (Relatives of rice have microtrichomes of 2 cells which secrete salt.)

(v) cotton fibres: it may be advantageous to increase the number of cotton fibres per boll, and at the same time decrease leaf trichomes to prevent insects hiding and prevent contamination of bolls.

(vi) ornamental plants: it may be preferred to decrease the number of hairs on hairy and glabrous varieties of a range of garden plants.

2. Trichomes on roots: manipulation of the number of root hairs may affect water and nutrient absorption (crop nutrient use efficiency) by the plants. Root hairs are also involved in anchoring the plant in the soil, particularly sandy soils, and allow better root penetration.

3. Seed mucilage: manipulation may lead to better seed germination in dry soils, due to maintenance of moisture

around the seed.

4. Seed dormancy: alteration of seed dormancy may allow quicker, or slower (if viviparous), germination of seeds after harvest. This may lead to faster cycling of crops.

5. Anthocyanin pigments: have a range of functions in the plant, and manipulation may alter pigmentation of seeds, leaves, flowers and fruit. Such manipulation may lead to:

(i) UV-B protection of plants, mainly in leaves, but anthocyanins are produced in plants under a wide range of stresses, including water stress, light stress, increased sugars. These stresses lead to decreased photosynthesis and susceptibility to photoxidation.

(ii) Altered flower and leaf colour in ornamentals and food crops, eg broccoli; altered fruit and seed colour in food, eg aubergines and grains (maize, rice, etc).

6. Condensed tannins, produced by the polymerisation of anthocyanin precursors, are found in many plants and are responsible in part for the taste characteristics of a range of fruits and vegetables, such as apple, kiwifruit, gooseberry, redcurrant and banana. Condensed tannins produce characteristic astringent properties in tea, coffee, wine, spices and fruit juices. Tannins also have important effects in animal feedstuffs. In monogastric animals, such as pigs and chickens, tannins limit the use of potential feedstuffs such as faba beans and sorghum. In ruminants, moderate levels of tannins are beneficial and may improve retention of dietary nitrogen, but higher levels reduce the nutritive value of foliage and feedstuffs. Manipulation of *TTG1* may alter the levels of condensed tannins in these plants.

7. Stomata on hypocotyls: increases in the number stomata may result in faster seedling growth under ideal conditions, such as optimum water and CO_2 availability.

The present invention provides a nucleic acid

isolate encoding a polypeptide including the amino acid sequence shown in Figure 3 (SEQ ID No. 2) or homologues thereof, which may include the coding sequence shown in Figure 3 which is that of the *TTG1* gene of *Arabidopsis thaliana*, and/or other transcribed parts of the gene e.g. as shown in Figure 3 or Figure 5.

Nucleic acid according to the present invention may have the sequence of an *TTG1* gene of *Arabidopsis thaliana*, or be a mutant, variant, derivative or allele or a homologue of the sequence provided. Preferred mutants, variants, derivatives and alleles are those which encode a protein which retains a functional characteristic of the protein encoded by the wild-type gene, especially the ability to affect a physical characteristic of a plant, such as the phenotypic characteristics outlined above.

A mutant, variant, derivative or allele in accordance with the present invention may have the ability to affect a physical characteristic of a plant, particularly a phenotypic characteristic identified above. Thus, a mutant, variant, derivative or allele may decrease the amount of anthocyanins in the epidermis of leaves and stems compared with wild-type on expression in a plant, e.g. compared with the effect obtained using a gene sequence expressing the polynucleotide sequence of Figure 3.

Alternatively or in addition, a mutant, variant, derivative or allele increases or decreases the number of trichomes on different organs compared with wild-type on expression in a plant, e.g. compared with the effect obtained using a gene sequence expressing the polynucleotide sequence of Figure 3. Down-regulation of TTG1 activity may be achieved by mutant nucleic acids (e.g through co-suppression) or by mutant polypeptides, which may compete for receptors or other binding sites for TTG1, without triggering appropriate effects.

Comparison of effect on the increase or decrease of

trichomas or other characteristics may be performed in *Arabidopsis thaliana*, although nucleic acid according to the present invention may be used in the production of a wide variety of plants and for influencing a phenotypic characteristic thereof.

Changes to a sequence, to produce a mutant, variant or derivative, may be by one or more of addition, insertion, deletion or substitution of one or more nucleotides in the nucleic acid, leading to the addition, insertion, deletion or substitution of one or more amino acids in the encoded polypeptide. Further, it may lead to the creation of stop codons resulting truncated polypeptide; removal of stop codons resulting in extended polypeptides; or a frameshift resulting in a polypeptide lacking *TTG1* function. Of course, changes to the nucleic acid which make no difference to the encoded amino acid sequence are included.

A preferred nucleic acid sequence for an *TTG1* gene is shown as the coding sequence within Figure 3/SEQ ID No. 1, alongside the predicted amino acid sequence of a polypeptide according to the present invention which has TTG1 function (SEQ ID No. 2).

Particular mutant alleles of the nucleic acid according to the present invention include:

a) ttg1.10 (SEQ ID No. 3) which contains a point mutation (G to A) in the 5′ untranslated part of the TTG1 sequence (see Fig 3);

b) ttg1.19 (SEQ ID No. 4) which results in the introduction of a stop codon at codon 183;

c) ttg1.1 (SEQ ID No. 5 - formerly designated ttg1.21) which results in the introduction of a stop codon at codon 317;

d) ttg1.20 (SEQ ID No. 6) which contains a point mutation (S to C) at codon 30, plus introduction of a stop codon at codon 310;

e) ttg1.9 (SEQ ID No.7) which contains a point mutation (S to F) at codon 282.

f) ttg1.15, ttg1.16, ttg1.17, ttg1.18 which all result in a stop codon at codon 310 (via a substitution of 2 different bases - TCGGCT to TAGACT - this sequence is designated SEQ ID No. 13).

Mutant alleles b) to f) contain point mutations that result in changes to the protein product. Mutations in the 5'untranslated leader sequences may affect the translation of the RNA.

Interestingly the present inventors have established that these mutations can lead to quite different phenotypes for the plants expressing them. For instance ttg1.9 has more anthocyanin present than ttg1.1 (reference allele) while ttg1.10 has more trichomes, different seed mucilage and less anthocyanin than ttg1.9 (see Larkin et al, Plant Cell 6, 1065-1076 for an analysis).

Thus it is clear that, given the sequence information disclosed herein, it will be possible for the skilled person, if desried, to generate ttg1 mutant having some (but not all) of the pleiotropic effects of the wild-type/ecotype TTG1 gene.

It will be appreciated by the skilled person that the above exemplified point mutations may be present individually or in combination with other point mutations. In other words, a mutant, allele, variant or derivative amino acid sequence in accordance with the present invention may include within the sequence shown in Figure 3, a single amino acid change with respect to the sequence shown in Figure 3, or 2, 3, 4, 5, 6, 7, 8, or 9 changes, about 10, 15, 20, 30, 40 or 50 changes, or greater than about 50, 60, 70, 80 or 90 changes. In addition to one or more changes within the amino acid sequence shown in Figure 3, a mutant, allele, variant or derivative amino acid sequence may include additional amino acids at the C-terminus and/or N-terminus.

A sequence related to a sequence specifically disclosed herein shares homology with that sequence.

Homology may be at the nucleotide sequence and/or amino acid sequence level. Preferably, the nucleic acid and/or amino acid sequence shares homology with the nucleotide sequence of Figure 3, or the amino acid sequence encoded thereby. Preferably the homology is at least about 50%, or 60%, or 70%, or 80% homology, most preferably at least about 90%, 95%, 96%, 97%, 98% or 99% homology.

As is well-understood, homology at the amino acid level is generally in terms of amino acid similarity or identity. Similarity allows for "conservative variation", i.e. substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine. Similarity may be as defined and determined by the TBLASTN program, of Altschul *et al.* (1990) *J. Mol. Biol.* **215**: 403-10, which is in standard use in the art, or, and this may be preferred, the standard program BestFit, which is part of the Wisconsin Package, Version 8, September 1994, (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA, Wisconsin 53711). BestFit makes an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman.

Homology may be over the full-length of the relevant sequence shown herein, or may more preferably be over a contiguous sequence of about or greater than about 20, 25, 30, 33, 40, 50, 67, 133, 167, 200, 233, 267, 300, 333 or more amino acids or codons, compared with the relevant amino acid sequence or nucleotide sequence as the case may be.

Also provided by an aspect of the present invention is nucleic acid including or consisting essentially of a sequence of nucleotides complementary to a nucleotide sequence with any sequence provided herein. Further,

there is provided nucleic acid including or consisting essentially of a sequence of nucleotides complementary to a nucleotide sequence hybridisable with any encoding sequence provided herein. Another way of looking at this would be for nucleic acid according to this aspect to be hybridisable with a nucleotide sequence complementary to any encoding sequence provided herein. Of course, DNA is generally double-stranded and blotting techniques such as Southern hybridisation are often performed following separation of the strands without a distinction being drawn between which of the strands is hybridising. Preferably the hybridisable nucleic acid or its complement encode a product able to influence a physical characteristic of a plant, particularly a phenotypic characteristic as described above. Preferred conditions for hybridisation are familiar to those skilled in the art, but are generally stringent enough for there to be positive hybridisation between the sequences of interest to the exclusion of other sequences.

For instance, screening may initially be carried out under conditions, which comprise a temperature of about 37°C or less, a formamide concentration of less than about 50%, and a moderate to low salt (e.g. Standard Saline Citrate ('SSC') = 0.15 M sodium chloride; 0.15 M sodium citrate; pH 7) concentration.

~~Alternatively, a temperature of about 50°C or less~~ and a high salt (e.g. 'SSPE'= 0.180 mM sodium chloride; 9 mM disodium hydrogen phosphate; 9 mM sodium dihydrogen phosphate; 1 mM sodium EDTA; pH 7.4). Preferably the screening is carried out at about 37°C, a formamide concentration of about 20%, and a salt concentration of about 5 X SSC, or a temperature of about 50°C and a salt concentration of about 2 X SSPE. These conditions will allow the identification of sequences which have a substantial degree of homology (similarity, identity) with the probe sequence, without requiring the perfect homology for the identification of a stable hybrid.

Suitable conditions include, e.g. for detection of sequences that are about 80-90% identical, hybridization overnight at 42°C in 0.25M Na_2HPO_4, pH 7.2, 6.5% SDS, 10% dextran sulfate and a final wash at 55°C in 0.1X SSC, 0.1% SDS. For detection of sequences that are greater than about 90% identical, suitable conditions include hybridization overnight at 65°C in 0.25M Na_2HPO_4, pH 7.2, 6.5% SDS, 10% dextran sulfate and a final wash at 60°C in 0.1X SSC, 0.1% SDS.

The nucleic acid, which may contain, for example, DNA encoding the amino acid sequence of Figure 3, as genomic or cDNA, may be in the form of a recombinant and preferably replicable vector, for example a plasmid, cosmid, phage or *Agrobacterium* binary vector. The nucleic acid may be under the control of an appropriate promoter or other regulatory elements for expression in a host cell such as a microbial, e.g. bacterial, or plant cell. In the case of genomic DNA, this may contain its own promoter or other regulatory elements and in the case of cDNA this may be under the control of an appropriate promoter or other regulatory elements for expression in the host cell.

A vector including nucleic acid according to the present invention need not include a promoter or other regulatory sequence, particularly if the vector is to be used to introduce the nucleic acid into cells for recombination into the genome.

Those skilled in the art are well able to construct vectors and design protocols for recombinant gene expression. Suitable vectors can be chosen or constructed, containing appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. For further details see, for example, *Molecular Cloning: a Laboratory Manual*: 2nd edition, Sambrook *et al*, 1989, Cold Spring Harbor Laboratory Press. Many known techniques and

protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in *Current Protocols in Molecular Biology*, Second Edition, Ausubel et al. eds., John Wiley & Sons, 1992. The disclosures of Sambrook et al. and Ausubel et al. are incorporated herein by reference. Specific procedures and vectors previously used with wide success upon plants are described by Bevan (Nucl. Acids Res. 12, 8711-8721 (1984)) and Guerineau and Mullineaux (1993) (Plant transformation and expression vectors. In: Plant Molecular Biology Labfax (Croy RRD ed) Oxford, BIOS Scientific Publishers, pp 121-148).

Selectable genetic markers may be used consisting of chimeric genes that confer selectable phenotypes such as resistance to antibiotics such as kanamycin, hygromycin, phosphinotricin, chlorsulfuron, methotrexate, gentamycin, spectinomycin, imidazolinones and glyphosate.

Nucleic acid molecules and vectors according to the present invention may be provided isolated and/or purified from their natural environment, in substantially pure or homogeneous form, or free or substantially free of nucleic acid or genes of the species of interest or origin other than the sequence encoding a polypeptide with the required function. Nucleic acid according to the present invention may include cDNA, RNA, genomic DNA and may be wholly or partially synthetic. The term "isolate" encompasses all these possibilities. Where a DNA sequence is specified, e.g. with reference to a figure, unless context requires otherwise, the RNA equivalent, with U substituted for T where it occurs, is encompassed.

When introducing a chosen gene construct into a cell, certain considerations must be taken into account, well known to those skilled in the art. The nucleic acid to be inserted should be assembled within a construct

which contains effective regulatory elements which will drive transcription. There must be available a method of transporting the construct into the cell. Once the construct is within the cell membrane, integration into the endogenous chromosomal material either will or will not occur. Finally, as far as plants are concerned the target cell type must be such that cells can be regenerated into whole plants.

Plants transformed with the DNA segment containing the sequence may be produced by standard techniques which are already known for the genetic manipulation of plants. Plants can be transformed with DNA using any suitable technology, such as a disarmed Ti-plasmid vector carried by Agrobacterium exploiting its natural gene transfer ability (EP-A-270355, EP-A-0116718, NAR 12(22) 8711 - 87215 1984), particle or microprojectile bombardment (US 5100792, EP-A-444882, EP-A-434616) microinjection (WO 92/09696, WO 94/00583, EP 331083, EP 175966, Green *et al.* (1987) *Plant Tissue and Cell Culture*, Academic Press), electroporation (EP 290395, WO 8706614 Gelvin Debeyser - see attached) other forms of direct DNA uptake (DE 4005152, WO 9012096, US 4684611), liposome mediated DNA uptake (e.g. Freeman *et al. Plant Cell Physiol.* 29: 1353 (1984)), or the vortexing method (e.g. Kindle, *PNAS U.S.A.* 87: 1228 (1990d) Physical methods for the transformation of plant cells are reviewed in Oard, 1991, *Biotech. Adv.* 9: 1-11.

Agrobacterium transformation is widely used by those skilled in the art to transform dicotyledonous species. Recently, there has been substantial progress towards the routine production of stable, fertile transgenic plants in almost all economically relevant monocot plants (Toriyama, et al. (1988) *Bio/Technology* 6, 1072-1074; Zhang, *et al.* (1988) *Plant Cell Rep.* 7, 379-384; Zhang, *et al.* (1988) *Theor Appl Genet* 76, 835-840; Shimamoto, *et al.* (1989) *Nature* 338, 274-276; Datta, *et al.* (1990) *Bio/Technology* 8, 736-740; Christou, *et al.* (1991)

Bio/Technology 9, 957-962; Peng, *et al.* (1991) International Rice Research Institute, Manila, Philippines 563-574; Cao, *et al.* (1992) *Plant Cell Rep.* 11, 585-591; Li, *et al.* (1993) *Plant Cell Rep.* 12, 250-255; Rathore, *et al.* (1993) *Plant Molecular Biology* 21, 871-884; Fromm, *et al.* (1990) *Bio/Technology* 8, 833-839; Gordon-Kamm, *et al.* (1990) *Plant Cell* 2, 603-618; D'Halluin, *et al.* (1992) *Plant Cell* 4, 1495-1505; Walters, *et al.* (1992) *Plant Molecular Biology* 18, 189-200; Koziel, *et al.* (1993) *Biotechnology* 11, 194-200; Vasil, I. K. (1994) *Plant Molecular Biology* 25, 925-937; Weeks, *et al.* (1993) *Plant Physiology* 102, 1077-1084; Somers, *et al.* (1992) *Bio/Technology* 10, 1589-1594; WO92/14828).. In particular, *Agrobacterium* mediated transformation is now emerging also as an highly efficient alternative transformation method in monocots (Hiei *et al.* (1994) *The Plant Journal* 6, 271-282).

The generation of fertile transgenic plants has been achieved in the cereals rice, maize, wheat, oat, and barley (reviewed in Shimamoto, K. (1994) *Current Opinion in Biotechnology* 5, 158-162.; Vasil, *et al.* (1992) *Bio/Technology* 10, 667-674; Vain et al., 1995, *Biotechnology Advances* 13 (4): 653-671; Vasil, 1996, *Nature Biotechnology* 14 page 702).

Microprojectile bombardment, electroporation and direct DNA uptake are preferred where Agrobacterium is inefficient or ineffective. Alternatively, a combination of different techniques may be employed to enhance the efficiency of the transformation process, eg bombardment with Agrobacterium coated microparticles (EP-A-486234) or microprojectile bombardment to induce wounding followed by co-cultivation with Agrobacterium (EP-A-486233).

Following transformation, a plant may be regenerated, e.g. from single cells, callus tissue or leaf discs, as is standard in the art. Almost any plant can be entirely regenerated from cells, tissues and organs of the plant. Available techniques are reviewed

in Vasil et al., *Cell Culture and Somatic Cell Genetics of Plants, Vol I, II and III, Laboratory Procedures and Their Applications*, Academic Press, 1984, and Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989.

The particular choice of a transformation technology will be determined by its efficiency to transform certain plant species as well as the experience and preference of the person practising the invention with a particular methodology of choice. It will be apparent to the skilled person that the particular choice of a transformation system to introduce nucleic acid into plant cells is not essential to or a limitation of the invention, nor is the choice of technique for plant regeneration.

A *TTG1* gene and modified versions thereof (alleles, mutants, variants and derivatives thereof), may be used to affect a physical characteristic, such as hairs on roots and aerial parts of plants and anthocyanin pigments characteristics, in plants. For this purpose nucleic acid such as a vector as described herein may be used for the production of a transgenic plant. Such a plant may possess an altered phenotype as described above compared with wild-type (that is to say a plant that is wild-type for *TTG1* or the relevant homologue thereof).

The invention further encompasses a host cell transformed with nucleic acid or a vector according to the present invention, especially a plant or a microbial cell. Thus, a host cell, such as a plant cell, including heterologous nucleic acid according to the present invention is provided. Within the cell, the nucleic acid may be incorporated within the chromosome. There may be more than one heterologous nucleotide sequence per haploid genome.

Also according to the invention there is provided a plant cell having incorporated into its genome nucleic acid, particularly heterologous nucleic acid, as provided by the present invention, under operative control of a

regulatory sequence for control of expression. The coding sequence may be operably linked to one or more regulatory sequences which may be heterologous or foreign to the gene, such as not naturally associated with the gene for its expression. The nucleic acid according to the invention may be placed under the control of an externally inducible gene promoter to place expression under the control of the user.

A suitable inducible promoter is the GST-II-27 gene promoter which has been shown to be induced by certain chemical compounds which can be applied to growing plants. The promoter is functional in both monocotyledons and dicotyledons. It can therefore be used to control gene expression in a variety of genetically modified plants, including field crops such as canola, sunflower, tobacco, sugarbeet, cotton; cereals such as wheat, barley, rice, maize, sorghum; fruit such as tomatoes, mangoes, peaches, apples, pears, strawberries, bananas, and melons; and vegetables such as carrot, lettuce, cabbage and onion. The GST-II-27 promoter is also suitable for use in a variety of tissues, including roots, leaves, stems and reproductive tissues.

A further aspect of the present invention provides a method of making such a plant cell involving introduction of nucleic acid or a suitable vector including the sequence of nucleotides into a plant cell and causing or allowing recombination between the vector and the plant cell genome to introduce the sequence of nucleotides into the genome. The invention extends to plant cells containing nucleic acid according to the invention as a result of introduction of the nucleic acid into an ancestor cell.

The term "heterologous" may be used to indicate that the gene/sequence of nucleotides in question have been introduced into said cells of the plant or an ancestor thereof, using genetic engineering, i.e. by human

intervention. A transgenic plant cell, i.e. transgenic for the nucleic acid in question, may be provided. The transgene may be on an extra-genomic vector or incorporated, preferably stably, into the genome. A heterologous gene may replace an endogenous equivalent gene, i.e. one which normally performs the same or a similar function, or the inserted sequence may be additional to the endogenous gene or other sequence. An advantage of introduction of a heterologous gene is the ability to place expression of a sequence under the control of a promoter of choice, in order to be able to influence expression according to preference. Furthermore, mutants, variants and derivatives of the wild-type gene, e.g. with higher or lower activity than wild-type, may be used in place of the endogenous gene. Nucleic acid heterologous, or exogenous or foreign, to a plant cell may be non-naturally occurring in cells of that type, variety or species. Thus, nucleic acid may include a coding sequence of or derived from a particular type of plant cell or species or variety of plant, placed within the context of a plant cell of a different type or species or variety of plant. A further possibility is for a nucleic acid sequence to be placed within a cell in which it or a homologue is found naturally, but wherein the nucleic acid sequence is linked and/or adjacent to nucleic acid which does not occur naturally within the cell, or cells of that type or species or variety of plant, such as operably linked to one or more regulatory sequences, such as a promoter sequence, for control of expression. A sequence within a plant or other host cell may be identifiably heterologous, exogenous or foreign.

Plants which include a plant cell according to the invention are also provided, along with any part or propagule thereof, seed, selfed or hybrid progeny and descendants. A plant according to the present invention may be one which does not breed true in one or more properties. Plant varieties may be excluded,

particularly registrable plant varieties according to Plant Breeders' Rights. It is noted that a plant need not be considered a "plant variety" simply because it contains stably within its genome a transgene, introduced into a cell of the plant or an ancestor thereof.

In addition to a plant, the present invention provides any clone of such a plant, seed, selfed or hybrid progeny and descendants, and any part of any of these, such as cuttings, seed. The invention provides any plant propagule, that is any part which may be used in reproduction or propagation, sexual or asexual, including cuttings, seed and so on. Also encompassed by the invention is a plant which is a sexually or asexually propagated off-spring, clone or descendant of such a plant, or any part or propagule of said plant, off-spring, clone or descendant.

The invention further provides a method of influencing or affecting a physical characteristic e.g. hairs on roots and aerial parts of plants and/or the presence or absence of anthocyanin pigment, including causing or allowing expression of a heterologous nucleic acid sequence as discussed within cells of the plant.

The invention further provides a method including expression from nucleic acid encoding the amino acid sequence of Figure 3, or a mutant, variant, allele or derivative of the sequence, within cells of a plant (thereby producing an encoded polypeptide), following an earlier step of introduction of the nucleic acid into a cell of the plant or an ancestor thereof. Such a method may influence or affect a phenotypic characteristic of the plant, such as those mentioned above. This may be used in combination with any other gene, such as transgenes involved in any other phenotypic trait or desirable property.

The present invention also encompasses the expression product of the nucleic acid sequence disclosed and methods of making the expression product by

expression from nucleic acid encoding therefor under appropriate conditions, which may be in appropriate host cells. Following expression, the product may be isolated from the expression system and may be used as desired, for instance in formulation of a composition including at least one additional component.

Purified TTG1 protein, or a variant thereof, e.g. produced recombinantly by expression from encoding nucleic acid therefor, may be used to raise antibodies employing techniques which are standard in the art. Antibodies and polypeptides comprising antigen-binding fragments of antibodies may be used in identifying homologues from other plant species.

Methods of producing antibodies include immunising a mammal with the protein or a fragment thereof. Antibodies may be obtained from immunised animals using any of a variety of techniques known in the art, and might be screened, preferably using binding of antibody to antigen of interest. As an alternative or supplement to immunising a mammal, antibodies with appropriate binding specificity may be obtained from a recombinantly produced library of expressed immunoglobulin variable domains, e.g. using lambda bacteriophage or filamentous bacteriophage which display functional immunoglobulin binding domains on their surfaces; for instance see WO92/01047.

A further aspect of the present invention provides a method of identifying and cloning *TTG1* homologues from plant species other than *Arabidopsis thaliana* which method employs a nucleotide sequence obtainable from that shown in Figure 3. Such a method may include the steps of preparing nucleic acid from plant cells under test, providing a nucleic acid molecule having a nucleotide sequence shown in Figure 3 or complementary to a nucleic acid sequence shown in Figure 3, contacting nucleic acid in said preparation with said nucleic acid molecule under conditions for hybridization of said nucleic acid

molecule to any said gene or homologue in said preparation, and identifying said gene or homologue if present by its hybridization with said nucleic acid molecule.

Sequences derived from these may themselves be used in identifying and in cloning other sequences. The nucleotide sequence information provided herein, or any part thereof, may be used in a data-base search to find homologous sequences, expression products of which can be tested for ability to influence characteristics described above. These may have *TTG1* function or the ability to modify characteristics including hairs on roots and aerial parts of plants and anthocyanin pigments. Alternatively, nucleic acid libraries may be screened using techniques well known to those skilled in the art and homologous sequences thereby identified then tested for requisite functionality.

Further, nucleotide sequences obtained from that shown in figure 3 may be used to isolate *TTG1* homologous from other species of plants by techniques such as hybridization and polymerase chain reaction (PCR).

PCR techniques for the amplification/identification of nucleic acid are described in US Patent No. 4,683,195.

The nucleic acid sequence provided herein readily allows the skilled person to design PCR primers.

Such oligonucleotide probes or primers, as well as the full-length sequence (and mutants, alleles, variants and derivatives) are also useful in identifying homologous sequences. Further, the present invention also extends to oligonucleotide probes or primers for amplification and/or identification which are obtainable by use of the sequence shown in Figure 3, optionally by selecting regions which are conserved with other sequences e.g. from the prior art. Alternatively it may be desirable to generate more specific primers by selecting regions of the TTG1 which are not homologous to other proteins such as an11.

In some preferred embodiments, oligonucleotides according to the present invention that are fragments of the sequence shown in Figure 3, or any mutant, allele, variant, or derivatives thereof, are at least 10 nucleotides in length, more preferably at least 15 nucleotides in length, more preferably at least 20 nucleotides in length.

Such fragments themselves individually represent aspects of the present invention.

Techniques corresponding to those above may also be used for ascertaining the genotype of mutant plants having altered phenotypes corresponding to TTG1 activities (e.g which lack trichomes or anthocyanin) i.e. the probes and primers of the present invention can be used for diagnosing mutations in such plants, or as markers for these traits.

As described above, the present invention also extends to nucleic acid encoding an *TTG1* homologue obtained using a nucleotide sequence derived from that shown in Figure 3.

In certain embodiments, nucleic acid according to the present invention encodes a polypeptide which has homology with all or part of the amino acid sequence shown in Figure 3, in the terms discussed already above (e.g. for length), which homology is greater over the length of the relevant part (i.e. fragment) (the relevant part being greater than 110 amino acids in length, preferably greater than 200 amino acids and even more preferably greater than 300 amino acids in length) than the homology shared between a respective part of the amino acid sequence of Figure 3 and the EST sequence, and may be greater than about 5% greater, more preferably greater than about 10% greater, more preferably greater than about 20% greater, and more preferably greater than about 30% greater.

Similarly, nucleic acid according to certain embodiments of the present invention may have homology

with all or part of the nucleotide sequence shown in Figure 3, in the terms discussed already above (e.g. for length), which homology is greater over the length of the relevant part (i.e. fragment) (the relevant part being greater than 350 nucleotide in length, preferably greater than 400 and even more preferably greater than 500 nucleotide in length) than the homology shared between a respective part of the nucleotide sequence of Figure 3 and may be greater than about 5% greater, more preferably greater than about 10% greater, more preferably greater than about 20% greater, and more preferably greater than about 30% greater. Thus, to exemplify with reference to one embodiment, nucleic acid may be provided in accordance with the present invention wherein the nucleotide sequence includes a contiguous sequence of about 350 nucleotides which has greater homology with a contiguous sequence of 350 nucleotides within the nucleotide sequence of Figure 3 than any contiguous sequence of 331 nucleotides of an EST sequence, preferably greater than about 5% greater homology, and so on.

The provision of sequence information for the *TTG1* gene of *Arabidopsis thaliana* enables the obtention of homologous sequences from other plant species. In particular, it should be possible to easily isolate *TTG1* homologues from related, commercially important Brassica species (e.g. *Brassica nigra, Brassica napus*, *Brassica campestris* and *Brassica oleracea*). Examples of homologues from *Matthiola incana* (ten week stock), *Nicotiana tobaccum* var Samsum (tobacco) and *Gossypium hirsutum* cv. Siokva 1-4 (cotton) are disclosed in the Examples below.

Thus, included within the scope of the present invention are nucleic acid molecules which encode amino acid sequences which are homologues of TTG1 of *Arabidopsis thaliana*. Homology may be at the nucleotide sequence and/or amino acid sequence level, as has already

been discussed above. A homologue from a species other than *Arabidopsis thaliana* encodes a product which causes a phenotype similar to that caused by the Arabidopsis thaliana *TTG1* gene, generally including the ability to influence a phenotypic characteristic, particularly a phenotypic characteristic as described above. In addition, mutants, derivatives or alleles of these genes may alter such characteristics compared with wild-type.

TTG1 gene homologues may also be identified from economically important monocotyledonous crop plants such as rice and maize. Although genes encoding the same protein in monocotyledonous and dicotyledonous plants show relatively little homology at the nucleotide level, amino acid sequences are conserved. Therefore it is possible to use public sequence databases to identify *Arabidopsis*, rice or maize cDNA clone sequences that were obtained in random sequencing programmes and share homology to the gene of interest, as has been done for flowering time genes isolated from Arabidopsis (e.g. CO; WO 96/14414).

Nucleic acid according to the invention may be used to modify the characteristics of a plant. This may be achieved by modification of expression of the nucleic acid according to the present invention or by interfering with the normal function of the protein encoded by the nucleic acid according to the present invention. For example, nucleic acid according to the present invention may be used to increase or decrease the number of trichomes on different organs to enhance their effectiveness. Further, it may be used to alter the pigmentation of seeds, leaves, flowers and fruit for UV protection and/or colour for presentation reasons or for ornamental plants. This may involve use of anti-sense or sense regulation, discussed further below.

As noted above, other physical characteristics of plants may be affected by means of expression from nucleic acid according to the present invention.

Nucleic acid according to the invention, such as an *TTG1* gene or homologue, may be placed under the control of an externally inducible gene promoter to place the timing of altering the characteristics of the plant under the control of the user. An advantage of introduction of a heterologous gene into a plant cell, particularly when the cell is comprised in a plant, is the ability to place expression of the gene under the control of a promoter of choice, in order to be able to influence gene expression, and therefore characteristic modification, according to preference. Furthermore, mutants and derivatives of the wild-type gene, eg with higher or lower activity than wild-type, may be used in place of the endogenous gene.

In the present invention, over-expression may be achieved by introduction of the nucleotide sequence in a sense orientation. Thus, the present invention provides a method of influencing a physical e.g. a phenotypic characteristic described above such as, an increase or decrease in trichomes, characteristic of a plant, the method including causing or allowing expression of the product (polypeptide or nucleic acid transcript) encoded by heterologous nucleic acid according to the invention from that nucleic acid within cells of the plant.

Down-regulation of expression of a target gene may be achieved using anti-sense technology or "sense regulation" ("co-suppression").

In using anti-sense genes or partial gene sequences to down-regulate gene expression, a nucleotide sequence is placed under the control of a promoter in a "reverse orientation" such that transcription yields RNA which is complementary to normal mRNA transcribed from the "sense" strand of the target gene. See, for example, Smith *et al*, (1988) *Nature* 334, 724-726; Zhang *et al*, (1992) *The Plant Cell* 4, 1575-1588, English *et al*., (1996) *The Plant Cell* 8, 179-188. Antisense technology is also reviewed in Bourque, (1995), *Plant Science* 105, 125-149, and Flavell, (1994) *PNAS USA* 91, 3490-3496.

An alternative is to use a copy of all or part of the target gene inserted in sense, that is the same, orientation as the target gene, to achieve reduction in expression of the target gene by co-suppression. See, for example, van der Krol *et al.*, (1990) *The Plant Cell* 2, 291-299; Napoli *et al.*, (1990) *The Plant Cell* 2, 279-289; Zhang *et al.*, (1992) *The Plant Cell* 4, 1575-1588, and US-A-5,231,020.

The complete sequence corresponding to the coding sequence (in reverse orientation for anti-sense) need not be used. For example fragments of sufficient length may be used. It is a routine matter for the person skilled in the art to screen fragments of various sizes and from various parts of the coding sequence to optimise the level of anti-sense inhibition. It may be advantageous to include the initiating methionine ATG codon, and perhaps one or more nucleotides upstream of the initiating codon. A further possibility is to target a conserved sequence of a gene, e.g. a sequence that is characteristic of one or more genes, such as a regulatory sequence.

The sequence employed may be about 500 nucleotides or less, possibly about 400 nucleotides, about 300 nucleotides, about 200 nucleotides, or about 100 nucleotides. It may be possible to use oligonucleotides of much shorter lengths, 14-23 nucleotides, although longer fragments, and generally even longer than about 500 nucleotides are preferable where possible, such as longer than about 600 nucleotides, than about 700 nucleotides, than about 800 nucleotides, than about 1000 nucleotides or more.

It may be preferable that there is complete sequence identity in the sequence used for down-regulation of expression of a target sequence, and the target sequence, though total complementarity or similarity of sequence is not essential. One or more nucleotides may differ in the sequence used from the target gene. Thus, a sequence

employed in a down-regulation of gene expression in accordance with the present invention may be a wild-type sequence (e.g. gene) selected from those available, or a mutant, derivative, variant or allele, by way of insertion, addition, deletion or substitution of one or more nucleotides, of such a sequence. The sequence need not include an open reading frame or specify an RNA that would be translatable. It may be preferred for there to be sufficient homology for the respective anti-sense and sense RNA molecules to hybridise. There may be down regulation of gene expression even where there is about 5%, 10%, 15% or 20% or more mismatch between the sequence used and the target gene.

Generally, the transcribed nucleic acid may represent a fragment of an *TTG1* gene, such as including a nucleotide sequence shown in Figure 3, or the complement thereof, or may be a mutant, derivative, variant or allele thereof, in similar terms as discussed above in relation to alterations being made to an *TTG1* coding sequence and the homology of the altered sequence. The homology may be sufficient for the transcribed anti-sense RNA to hybridise with nucleic acid within cells of the plant, though irrespective of whether hybridisation takes place the desired effect is down-regulation of gene expression.

Thus, the present invention also provides a method of influencing a characteristic of a plant such as any one of those described above, the method including causing or allowing anti-sense transcription from heterologous nucleic acid according to the invention within cells of the plant.

The present invention further provides the use of the nucleotide sequence of Figure 3 or a fragment, mutant, derivative, allele, variant or homologue thereof for down-regulation of gene expression, particularly down-regulation of expression of an *TTG1* gene or homologue thereof, preferably in order to influence a

physical characteristic of a plant, especially a phenotypic characteristic such as an increase or decrease of trichomes on different organs and/or and increase or decrease in anthocyanin pigments.

When additional copies of the target gene are inserted in sense, that is the same, orientation as the target gene, a range of phenotypes is produced which includes individuals where over-expression occurs and some where under-expression of protein from the target gene occurs. When the inserted gene is only part of the endogenous gene the number of under-expressing individuals in the transgenic population increases. The mechanism by which sense regulation occurs, particularly down-regulation, is not well-understood. However, this technique is also well-reported in scientific and patent literature and is used routinely for gene control. See, for example, van der Krol *et al.*, (1990) *The Plant Cell 2*, 291-229; Napoli *et al.*, (1990) The Plant Cell *2*, 279-289; Zhang et al, 1992 *The Plant Cell 4*, 1575-1588.

Again, fragments, mutants and so on may be used in similar terms as described above for use in anti-sense regulation.

Further options for down regulation of gene expression include the use of ribozymes, e.g. hammerhead ribozymes, which can catalyse the site-specific cleavage of RNA, such as mRNA (see e.g. Jaeger (1997) "The new world of ribozymes" Curr Opin Struct Biol 7:324-335, or Gibson & Shillitoe (1997) "Ribozymes: their functions and strategies form their use" Mol Biotechnol 7: 242-251.).

Another option is the use of nucleic acids enconding non-functional or partially functional mutant proteins (e.g. encoded by the mutant alleles of the present invention, or as produced by mutagenesis) which, when expressed in a plant, may compete with functional TTG1 proteins for e.g. receptors or other binding partners thereby reducing the effectiveness of those proteins.

Thus, the present invention also provides a method

of influencing a phenotypic characteristic of a plant, the method including causing or allowing expression from nucleic acid according to the invention within cells of the plant. This may be used to suppress activity of a product with ability to influence a phenotypic characteristic as described above. Here the activity of the product is preferably suppressed as a result of under-expression within the plant cells.

Also embraced within the present invention are untranscribed parts of the TTG1 gene. Thus in a further aspect of the present invention there is disclosed a nucleic acid molecule encoding the promoter of the TTG1 gene. Owing to the widespread presence of the TTG1 transcript in the plant it is believed that this promoter is constitutive or essentially constitutive, and thus may have utility in producing constructs for the expression of genes in plants. Variant promoters having promoter activity are also embraced by the present invention.

To find homologous promoters, or the minimal elements or motifs responsible for promoter activity, restriction enzyme or nucleases may be used to digest a nucleic acid molecule comprising the 5′ region of Seq ID No 1, or mutagenesis may be employed, followed by an appropriate assay (for example using a reporter gene such as luciferase operably linked to the restricted sequence). Methods for promoter identification may employed without burden by those skilled in the art in the light of the sequence data disclosed herein. Once characterised the promoters of the present invention may incorporated into vectors.

Aspects and embodiments of the present invention will now be illustrated, by way of example, with reference to the accompanying figures. Further aspects and embodiments will be apparent to those skilled in the art. All documents mentioned in this text are incorporated herein by reference.

Brief description of the drawings

Figure 1

This shows A) a large scaled map(not to scale) of the TTG1 region showing the relationship between probes used in the region. YAC end probes are indicated by boxes; and B) a fine scale map of the TTG1 region determined by RFLP mapping. End probes generated from YACs are indicated as boxes. The position of the cosmid g4556 is shown in relation to the genomic lambda isolated from this region. The recombination points are marked with a cross.

Figure 2

This shows a map of genomic clones used to complement *ttg1* mutants . It shows the genomic fragments that have been used to complement the *ttg1* mutant phenotype. Fragment(a) represents the genomic 13.8kb insert in pB8 with the EcoRI (E) and XbaI (X) restriction sites some of which are used to create the deletion. Fragment (b) represents pB8DX1, (c) represents pB8DE2 and (d) represents pB8DE3. Deletions are indicated by the dotted lines. Fragments (a) and (d) both gave transformed plants with trichomes and anthocyanin. Transformants of (b) and (c) lacked both trichomes and anthocyanin.

Figure 3A and 3B

This shows the sequence of *TTG1* locus (SEQ ID No. 1). The intron (coding sequence) is in italics and the predicted amino acid sequence (SEQ ID No. 2) is shown by the single letter code under the nucleotide sequence. Five identified mutations are shown (ttg1.10, ttg1.19, ttg1.20, ttg1.1, ttg1.9 - SEQ ID No's. 3 to 7 respectively). Sixty bases are shown per line.

Figure 4

This shows an alignment of TTG1, AN11 from petunia

and the partial sequences from Matthiola (Seq ID No 8) and tobacco (Seq ID Nos 9 and 10).

Figure 5

Shows the predicted cDNA sequence of *TTG1.* This corresponds to the region shown in capitals in Fig 3, plus a further 10 nucleotides which were subsequently mapped to the end of the transcript by primer extension studies.

Summary of sequence ID Nos

1: Full length TTG1 DNA sequence shown in Fig 3, including promoter region, full cDNA sequence, and coding sequence (which is aligned with the amino acid sequence).
2: TTG1 amino acid sequence in Fig 3 and 4.
3: DNA sequence ttg1.10
4: DNA sequence ttg1.19
5: DNA sequence ttg1.1
6: DNA sequence ttg1.20
7: DNA sequence ttg1.9
8: Partial amino acid sequence of Matthiola TTG1 homologue
9: Partial amino acid sequence of first tobacco TTG1 homologue.
10: Partial amino acid sequence of second tobacco TTG1 homologue.
11 and 12: Degenerate primers for cloning TTG1 homologues (see Examples below).
13: DNA sequence of ttg1.15, ttg1.16, ttg1.17 and ttg1.18.

Detailed description.

Molecular Mapping of *ttg1.1*

Recombinants between the *ttg1.1* and *MS1* loci generated in a cross between Landsberg *erecta* carrying *ttg1* and *ms1* and Ws ecotypes were analysed using RFLPs (restriction fragment length polymorphism) between these

parents with probes already mapped to this region by Nam et al (Nam et al (1989) Plant Cell 1 699-705.). Recombinants on the distal side of *ttg1.1* were selected from a cross of Landsberg *erecta* carrying the *ttg1.1* mutation and *ga3* and *ch5* and the RLD1 ecotype which was wildtype for these loci. End probes from YACs that had been mapped to the region between *ms1* and *ga3* (Schmidt et al (Schmidt et al (1997) Plant J 11 563-572)) were also utilized to map the location of *ttg1*. RFLPs generated by the cosmid g4556 could not be separated from *ttg1* with the exception of one recombinant called Dennis 19 (on the *ms1* side of *ttg1*) suggesting that g4556 was very close to the mutation in *ttg1.1*. A YAC EG20H2 that hybridised to g4556 and the cosmid were used to isolate overlapping genomic lambda clones. The lambda clones were ordered using restriction mapping and hybridization techniques and then used as probes for RFLPs amongst the recombinants. Figure 1 shows the large scale and fine scale maps from the *TTG1* region that were derived from the analysis of the recombinants on both sides of the *ttg1.1* mutation. In Figure 1B the position of several recombination events between *ms1* and *ttg1* have been indicated. On the distal side of *ttg1* no nearby recombination events could be mapped due to lack of RFLPs between the ecotypes used and lack of success in isolating clones from this region from three different libraries in lambda or cosmid vectors using a variety of probes.

<u>Complementation of the *ttg1* mutation</u>

The complete genomic inserts from overlapping lambda clones 1.1A, 3.1A, 8 and X6 marked in Figure 1B were subcloned into pBinNOT using the flanking NotI restriction sites in the lambda vector, giving a pB series of binary vectors. These were transferred into *Agrobacterium tumefaciens* strain AGL1 (Lazo et al (1991) Biotechnology 9 963-967) and then used to transform

Arabidopsis ecotype Landsberg *erecta* carrying the *ttg1.1* mutation by co-cultivation with root explants (Valvekens et al (1988) Proc Natl Acad Sci USA 85 5536-5540). Only a small number of kanamycin-resistant transformants were obtained, but one plant from pB8, derived from lambda 8, had trichomes but failed to set seed. Transformants from pB1.1A did not have trichomes. Other kanamycin-resistant shootlets appeared to be escapes due to prolonged exposure to kanamycin in the callus stage.

Several deletions were made of pB8 utilizing restriction sites within the genomic sequence and the polylinker of the vector. These deletion constructs (shown in Figure 2) were used to transform Arabidopsis ecotype Columbia carrying the *ttg1.9* mutation via vacuum infiltration (Bechtold et al (1993) Compt Rend Acad Sci III-Life Sci 316 1194-1199). Sixty transformants from pB8DE3 (indicated as construct d) produced trichomes, although one transformant from pB8DE3 showed kanamycin resistance but had no trichomes. The transformants bearing trichomes exhibited other wild-type characteristics of brown seed, seed mucilage, purple colouring of the plant and normal root hair numbers indicating that the other *ttg1* mutant phenotypes had also been complemented. Thirty-six kanamycin-resistant transformants from pB8DE2 (construct c) and 19 transformants from pB8DX1 (construct b) failed not produce trichomes, suggesting that TTG1 was located in the regions deleted in these constructs.

The positional cloning of the *TTG1* locus has provided information about the order of, and distances between, a number of RFLP markers which may be used to isolate nearby genes. This information is complementary to the data given in the physical maps of the region (Schmidt et al., 1997 Plant J 11: 563-572; Thorlby et al., 1997 Plant J 12, 471-479). Although the present inventors analysed about 400 recombinants within a 14 map unit region, they were still unable to find breakpoints

very close to the *TTG1* locus. This suggests that recombination rates are reduced close to this gene. Recombination frequencies are known to vary along chromosomes in many species (Lichten and Goldman, 1995 Annu Rev Genet 29: 445-476 for a review).

Subcloning of pB8DE3 and sequencing

Restriction fragments of the insert in pB8DE3 were ligated into pBluescript, sequenced using fluorescent dideoxynucleotides and the sequences compiled and analysed using the GCG package. Sequence analysis with Genmark and Netgene revealed two possible genes in the 5777 bp insert in pB8DE3.

One of these genes revealed no consistent homology to any known protein, was not similar to any EST clone in data-bases, and did not appear to code a long ORF.

The other predicted gene corresponded to an Arabidopsis EST (F20055, F20056), indicating that the gene was functional and expressed. The predicted protein sequence of 341 amino acids shows sequence similarity (about 45 %) to a large and diverse group of proteins with WD40 repeat motifs. There are seven WD40 repeats with a short N-terminal region. The first two repeats contain a proline-rich region - the second repeat having 8/23 amino acids which are proline. Three possible TATA boxes have been identified 133, 189 and 216 bases upstream of the predicted start of translation. Comparison of the genomic and EST sequences indicated the presence of a single intron 3′ of the termination codon. The sequence including the promoter region is shown in Figure 3. Primer extension experiments indicated that the start of transcription is 109 bases 5′ of the start of translation (i.e. 23 bases from a TATA box).

Sequence analysis of *ttg1* mutants

The present inventors examined the nucleotide sequence of this region of a number of *ttg1* mutant

alleles to determine whether the gene encoding the WD40 repeat protein was likely to be the *TTG1* locus. PCR products from the region were generated with primers designed to give overlapping fragments of about 700 bp. These PCR products were obtained from the *ttg1* mutants *ttg1.9, ttg1.10, ttg1.19* and *ttg1.1* and from their parental wild-type alleles using genomic DNA as the template. The PCR products were gel-purified and sequenced using both oligonucleotides designed as primers for PCR.

Four of the mutant alleles contained point mutations that would result in changes to the protein product (Figure 3). Point mutations in *ttg1.19* and *ttg1.1* resulted in the introduction of stop codons at codons 183 and 317, respectively.

A point mutation in *ttg1.20* resulted in the change of serine to cysteine at codon 30. This allele, plus also alleles *ttg1.15-18* contained a premature stop codon at position 310.

A mutation in *ttg1.9* resulted in the change of serine to phenylalanine at codon 282.

The mutant allele *ttg1.10* contained a point mutation (G to A) in the 5′ untranslated leader sequence, which may affect the translation of the RNA. These changes in the gene encoding the WD40 repeat protein confirm its identity as *TTG1*.

Several of the mutants have a lower level of transcript detected by northern analysis and compared with message levels from the wildtype parent (results not shown). The reduction in message level in mutants such as *ttg1-9* could be due to nonsense-mediated mRNA decay which has been shown to occur in plants as well as in other organisms (Dickey et al., 1994 Plant Cell 6, 1171-1176; van Hoof and Green, 1996 Plant J 10: 415-424).

In summary the phenotypes (all were ttg1-like) and mutations are as follows:

Nasc No	Parent	Nasc Decription	Allele	Mutation
N300	An-1	pale, branched	ttg1.15	s310 -> *
N319	En-2	dwarf	ttg1.16	s310 -> *
N339	En-2	pale	ttg1.17	s310 -> *
N372	En-2	upright rosette	ttg1.16	s310 -> *
N406	En-2		ttg1.19	w183 -> *
N420	En-2	early flowering	ttg1.20	s30 -> c s310 -> *
N447	En-2	dwarf	ttg1.1	(as ttg1.1)

Effect of TTG1 on stomata of Arabidopsis

TTG1 effect on stomata appears to be analogous to control of root hairs (Berger et al, 1998, Dev Biol 194: 226-234).

The table below shows a comparison between the stomatal numbers on hypocotyls in ttg1 mutants compared to wild-type in air and at elevated CO_2 concentrations.

	Air	CO_2
Landsberg erecta:	23.2±3.2	8.5±1.9
ttg1 mutant:	22.8±1.4	16.0±1.6

Analysis of the expression of the *TTG1* gene

To determine the length of the *TTG1* transcript and to see if the expression of the gene was confined to some organs, an RNA blot was hybridized with a *TTG1* probe. The resulting band was measured to be 1.35 kb in length and present in all organs tested (roots, rosette leaves, leaf buds, stems, cauline leaves, siliques, flowers, floral buds). A surprise result was that it was highly expressed in floral meristems where there are only a few trichomes on the sepals and no anthocyanin in the flower petals.

The high level of transcripts of the *TTG1* locus

suggests two possible points of regulation of the gene. The first is that the TTG1 protein is present in many tissues where it requires a partner for activation. Another possibility is that regulation of this gene occurs a posttranscriptional stage with protein only being present in those cells that require functional TTG1 protein.

Structure and function of the *TTG1* locus

The present inventors have identified the *TTG1* locus as a gene encoding a WD40 repeat protein by complementation of a *ttg1* mutant with genomic DNA and by sequencing the gene in several *ttg1* mutant alleles. Two of the mutant alleles contained stop codons that would result in the production of truncated proteins, and two others contained point mutations that would change serine residues, to a cysteine and phenylalanine residues. The TTG1 protein bears no resemblance to the maize *R* gene product which was able to complement the *ttg1* mutant phenotype in Arabidopsis and anthocyanin pigment in tobacco flowers (Lloyd et al (1992) Science 258 1773-1775). This suggests that TTG1 acts upstream of an Arabidopsis R homologue in the pathways leading to trichome differentiation and anthocyanin synthesis. The involvement of other WD40 proteins in signal transduction pathways (see below) suggests that *TTG1* is involved in a pathway, or pathways, regulating the expression or action of downstream transcription factors.

Computer modelling of TTG1

WD40 repeat proteins are involved in a number of different types of regulatory roles, such as signalling (eg. G_β subunit of heterotrimeric G proteins), cell cycle regulation (eg CDC20 and CDC4), transcriptional repression (eg yeast TUP1, *Drosophila* extra sex combs), vesicular trafficking (eg SEC13) and RNA processing (Neer et al (1994) Nature 371 297-300). The TTG1 protein shows

the highest sequence similarity to G_β subunits, which are the best characterised of the WD40 repeat proteins. The G_β subunit contains 7 repeats of the WD40 motif and has a structure resembling a seven-bladed propeller, based on its crystal structure with its partner G_γ (Sondek et al (1996) Nature 379 369-374). Each blade is composed of 4 β-sheets, and an N-terminal amphipathic α-helix interacts closely with the G_γ subunit which is required for correct folding and function of the G_β subunit. However, computer-aided modelling of TTG1 (in collaboration with N. Srinivasan and T.L. Blundell) suggests that proline-rich regions of the first two WD40 repeats may disrupt the folding of the β-sheets essential for the structure of each blade. In addition, amino acid residues identified as interacting with G_α and G_γ subunits are not well conserved in TTG1. This may suggest that TTG1 represents a separate class of WD40 repeat protein. Genes encoding several WD40 proteins, including G_β subunits and the COP1 protein, have been isolated from plants (Ma (1994) Plant Mol Biol 26 1611-1634), but none of these proteins closely resembles TTG1 other than an11.

Thus studies of other WD40 proteins make it probable that the TTG1 protein does not act directly as a transcription factor but binds to other proteins to promote the initiation of trichomes in leaves and stems. The TTG1 protein may act as part of a DNA binding complex to regulate transcription. However the amino acid sequence contains no recognizable nuclear localization signal from computer analysis, although a cryptic site might be present. Another possibility is that another protein is required to form a complex for nuclear import as is the case with AP3 and PI from Arabidopsis (McGonigle et al., 1996 Genes Dev 10, 1812-1821). Another possibility is that the TTG1 protein is only located in the cytoplasm and acts as part of a signal transduction pathway. GUS-TTG1 and TTG1-GFP fusion proteins appear to be cytoplasmically located.

There are several sequences of unknown function that show higher similarity to TTG1 than the G_β subunit from either plants or animals. One sequence is the result of a genomic sequencing project in Arabidopsis and has a 85% similarity to TTG1 and is located on chromosome 3. Gel blots hybridized and washed at low stringency show at least three bands in Arabidopsis and maize suggesting that TTG1 could belong to a class of proteins. Two *C.elegans genes* arrayed in tandem have greater similarity to TTG1 than any locus from yeast *Saccaromyces cerevici*. This is surprising as yeast is more closely related to Arabidopsis than C.elegans is. However if TTG1 plays a role in defining functions in epidermal cells, this function may also be required in other multicellular organisms.

Transcription factors like those from maize and *Antirrhinum* have been identified in *Petunia* where they regulate flower colour (reviewed in Mol et al., 1996). *JAF13* is similar to the *R* gene and *Delila* and *AN2* encodes a MYB factor like *C1* from maize. These are thought to act together in a similar way to R and C1 to positively regulate the anthocyanin pathway. *AN11* from Petunia controls anthocyanin biosynthesis in flowers (de Vetten et al., 1997 Genes Dev 11, 1422-1434) possibly by regulating *AN2*. This is contrary to the evidence that TTG1 might regulate MYC transcription factors, determined by overexpression of the maize *R* gene in Arabidopsis (Lloyd et al., 1992 Proc Nat Acad Sci USA 86: 7092-7096). The identification of two WD40 repeat proteins which regulate anthocyanins and, in the case of the TTG1 protein, many other pathways suggests that this class of protein may be involved in regulating developmental pathways in other organisms.

<u>Cloning of the Matthiola incana, Nicotiana tobaccum and cotton homologues</u>

Primers for degenerate PCR were designed by

comparison of the TTG1 sequence with the an11 sequence (see Fig 4). Primers were based on the sequence encoding amino acids 74-85, and 296-307.

Primer sequences were:

5′ (TL3) TTYGAICAYCCITAYCCICCIACIAARYTIATGTT (Seq ID No 11)
3′ (TL4) CATIGGRTCRATICCRTTIGGICCIGCIACIGTNGG (Seq ID No 12)

Amplification conditions were ascertained using a temperature gradient in a Robocycler (Stratagene). 1ng of genomic DNA template was amplified with 5pmol of each primer using Taq DNA polymerase (Qiagen) with the addition of QX (Qiagen) for stabilising the DNA-primer complex.

Once an annealing temperature of 42°C was established, genomic DNA from *Matthiola incana* (ten week stock). *Nicotiana tobaccum* var Samsum (tobacco) and *Gossypium hirsutum* cv. Siokva 1-4 (cotton) was used as a template in 50ul reactions with the temperature ramped between anncaling and extension to 15 degrees per minute.

The amplified bands were size-fractioned and extracted from a gel. The Matthiola DNA was polished with Klenow enzyme to make a blunt end ligated into the EcoRV site of pBluescript. Three independent constructs were sequenced. The gel purified PCR product was used as a probe on a Southern blot to verify that the product originated from Matthiola. At high stringency (65°C in 0.1XSSC+1% SDS) the PCR product cross hybridises to the TTG1 gene of Arabidopsis.

Tobacco sequences (tobacco 1 and 2) were obtained from two constructs (pTOB1 and pTOB2) using T-vectors bases on pBluescript. Each corresponds to one of the genomic sequences found in tobacco which is an allotetraploid species. The two sequences are 95%

identical at both the nucleotide and the amino acid level. pTOB1 hybridises to both tobacco genes but only weakly to TTG1 at high stringency. The cotton gene is currently being sequenced.

Garden blots reveal the presence of several similar sequences in Arabidopsis, tobacco, Petunia and *zea mays* when hybridised at 50°C in 5M NaCl and washed at 50°C in IXSSC+1% SDS.

Use of TTG1 anti-sense constructs

The insert in construct pTOB1 was removed using SacI and EcoRV and ligated into the SacI and SmaI sites in the pROK2 vector. This gives an antisense construct with the 35S promoter driving a transcript from the complementary strand of the TOB1-TTG1 gene. Constructs containing the TOB1 sequence may be placed into the Agrobacterium strain LBA4404 for transfer into tobacco plants.

General methods

General methods were perfomed in accordance with Sambrook et al (1989) discussed above.

Plant material

Ecotypes Landsberg *erecta*, Columbia, RLD1, Ws (Wassilewskija) were supplied by the Nottingham Arabidopsis Stock Centre. The Landsberg *erecta* line carrying *ms1* and *ttg1.1* was from . The line containing *ttg1.1 ga3 ch5* in Landsberg *erecta* background was a gift.

Growing plants

a) for crosses and seed in the 'Arabicon system', 3:2:1 soil vermiculite perlite was used with 16 hours light.
b) for material in trays, 12 hours light was used.
c) for transformation, peat-based soil with 5-10 plants in 4 inch pots in a glasshouse with supplementary lighting was used. Pots were covered in muslin, and seed mixed with sand was sprinkled on top. Plants were thinned

to 10 per pot at 2-3 weeks. Bolts were cut back once and allowed to reemerge for several days before infiltration. All three types of plants received a weekly feed of macro nutrients.

d) in culture 1/2 MS + 0.8% agar 16 hours light was used.

e) for root material, plants were grown on plates containing 1/2MS +1.2% phytogel in a nearly vertical position.

Searching the Kranz collection for more ttg1 alleles.

Candidates described as glabrous and having yellow seeds were grown and crossed to *ttg1.1* mutants. The F2 generation was examined for segregating phenotypes. All were examined for seed mucilage, anthocyanin in the plant and the testa and for leaf hairs.

DNA extractions

Plant material was treated as in Dellaporta et al.(1983) Plant Mol Biol Rep 1,4, 19-21, followed by CsCl banding to remove RNA and polysaccharides (Walker et al, 1997 Photosyn Res 54, 155-163) so as to be able to detect small band shifts on DNA gel blots.

Library screens

Genomic library in lambda DashII (Stratagene) from Landsberg *erecta* (Boyce et al, 1994 Plant Physiol 106, 1691) distributed by EEC-BRIDGE Arabidopsis DNA Stock Centre.

Construction of the pBINNOT vector

The pBIN19 vector was modified to contain a NotI site in the polylinker. To remove the original NotI site, pBIN19 was restricted with NotI, treated with Klenow and dNTPs to fill in the site, ligated in a large volume, restricted again with NotI and transfected into E.coli strain TG1. Plasmid DNA was isolated from resulting colonies to check that the original NotI site no longer

existed. The vector was restricted with XbaI and Asp718. Two annealed oligonucleotides (N1: GTACCGCGGCCGCAT AND N2: CTAGATGCGGCCGCG) containing a NotI site were ligated into the vector to reconstitute the XbaI and Asp718 sites. The ligated DNA was restricted with BamHI to remove parental molecules. In effect, the BamHI site in the polylinker of pBIN19 has been replaced with a unique NotI site and the altered vector called pBINNOT.

Plant transformation

Agrobacterium strain Agl1 (Lazo et al., 1991 Biotechnology 9, 963-967) was transformed with constructs based on pBINNOT vector by electroporation. Using vacuum infiltration (Bechtold et al, 1993 Compt Rend Acad Sci III Life Sci 316, 1194-1199). DNA containing genomic fragments were introduced into ttg1.9 mutant plants.

Claims

1. An isolated nucleic acid molecule encoding a polypeptide with *TTG1* function.

2. A nucleic acid as claimed in claim 1 wherein the polypeptide includes the amino acid sequence SEQ ID No. 2

3. A nucleic acid as claimed in claim 2 comprising a nucleotide sequence having SEQ ID NO 1 or the coding sequence shown therein

4. An isolated nucleic acid comprising a nucleic acid sequence which shares at least 50%; 60%; 70%; 80%; 90%; 95%; 96%; 97%; 98%; 99% sequence identity with the nucleic acid of claim 2 or claim 3.

5. A nucleic acid as claimed in claim 4 which is a mutant , variant, derivative of any one of the nucleic acid sequences of claim 2 or claim 3 by way of addition, insertion, deletion or substitution of one or more nucleotides in the nucleic acid.

6. A nucleic acid as claimed in claim 4 which is a TTG1 homologue from a species other than *Arabidopsis thaliana*

7. A nucleic acid as claimed in claim 6 which is a homologue from *nicotiana* or *matthiola*.

8. A nucleic acid as claimed in claim 7 which encodes a polypeptide comprising amino acid Seq ID No 8, 9 or 10.

9. A nucleic acid as claimed in claim 4 which is an allele of the nucleic acid sequences of claim 2 or claim 3.

10. A nucleic acid as claimed in any one of claims 4 to 9 which encodes a polypeptide with *TTG1* function.

11. A nucleic acid as claimed in any one of claims 1 to 3 or claim 10 wherein the TTG1 function comprises the ability to alter two or more of the following phenotypic characteristics of a plant into which said polypeptide is introduced: number of trichomes on the aerial parts of the plant; number of trichomes on the roots hairs; mucilage of the seeds; dormancy of the seeds; anthocyanin pigmentation; condensation of the tannins; number of stomata on hypocotyls.

12. A nucleic acid as claimed in claim 9 which comprises the nucleotide sequence of an allele selected from: ttg1.10, ttg1.19, ttg1.1, ttg1.20, ttg1.9.

13. A nucleic acid including or consisting essentially of a sequence of nucleotides complementary to the nucleotide sequence of a nucleic acid as claimed in any one of the preceding claims.

14. A method of identifying a homologue or allele as claimed in any one of claims 6 to 9 which method employs a nucleotide sequence obtainable from that shown in Seq ID No 1, said method including the steps of:
(i) preparing nucleic acid from plant cells under test,
(ii) providing a nucleic acid molecule which is a probe or primer having a nucleotide comprising all or part of a nucleotide sequence as claimed in claim 2 or claim 3, or complementary to that sequence
(iii) contacting nucleic acid in said preparation with said probe or primer under conditions for hybridization, and
(iv) identifying said gene or homologue if present by its hybridization with said nucleic acid molecule.

15. A method as claimed in claim 14 further comprising the step of testing the homologue or allele for TTG1 function.

16. A method as claimed in claim 14 or claim 15 wherein the plant cell nucleic acid is obtained from a plant species other than *Arabidopsis thaliana*.

17. An oligonucleotide for use as a nucleic acid probe or primer for use in the method as claimed in any one of claims 14 to 16, said oligonucleotide comprising:
(i) a nucleotide sequence encoding an amino acid sequence which is conserved between TTG1 and the polypeptide encoded by the nucleic acid of any one of claims 6 to 8; or
(ii) a nucleotide sequence which is complementary to said sequence.

18. An oligonucleotide as claimed in claim 17 comprising at least about 10; 15; 20; 25; 30 or 35 nucleotides in length.

19. A recombinant vector comprising the nucleic acid of any one of claims 1 to 13.

20. A vector as claimed in claim 19 wherein the nucleic acid is under the control of a promoter.

21. A vector as claimed in claim 20 wherein the promoter is an inducible promoter.

22. A vector as claimed in any one of claims 19 to 21 further comprising one or more of: a terminator sequence; a polyadenylation sequence; an enhancer sequence; a marker gene.

23. A host cell comprising the nucleic acid of any one

of claims 1 to 13.

24. A host cell transformed with nucleic acid of any one of claims 1 to 13 or a vector of any one of claims 19 to 22.

25. A host cell having incorporated into its genome heterologous nucleic acid as claimed in any one of claims 1 to 13.

26. A host cell as claimed in any one of claims 23 to 25 which is a plant cell.

27. A method of making the plant cell of claim 26, the method comprising the steps of:
(i) introducing a vector as claimed in any one of claims 19 to 22 into the plant cell,
(ii) causing or allowing recombination between the vector and the plant cell genome to introduce a nucleic acid as claimed in any one of claims 1 to 13 into the genome.

28. A plant which has been regenerated from the plant cell of claim 26.

29. A plant as claimed in claim 28 including the plant cell of claim 26.

30. A plant as claimed in claim 29 which is a clone; selfed or hybrid progeny, or other offspring or descendant of the plant of claim 28.

31. A cutting, part, or seed or other propagule of a plant as claimed in any one of claims 28 to 30.

32. A polypeptide expression product of any one of the nucleic acids of claims 1 to 11 which has TTG1 activity.

33. A polypeptide as claimed in claim 32 comprising the amino acid sequence SEQ ID No 2.

34. A method of making the polypeptide of claim 32 or claim 33 by causing or allowing expression from a nucleic acid encoding the polypeptide, following an earlier step of introduction of the nucleic acid into a cell of a plant or an ancestor thereof.

35. Use of a polypeptide of claim 32 or claim 33 or a nucleic acid of any of claims 1 to 13 to regulate the expression or action of a transcription factor.

36. Use of a polypeptide of claim 32 or claim 33 to raise an antibody.

37. An antibody having specific binding affinity for the polypeptide claimed in claim 32 or claim 33.

38. A polypeptide comprising the antigen-binding site of the antibody of claim 37.

39. A method of influencing or affecting a physical characteristic of a plant comprising causing or allowing expression of a heterologous nucleic acid sequence as claimed in any one of claims 1 to 13 within the cells of the plant, following an earlier step of introducing the nucleic acid into a cell of the plant or an ancestor thereof.

40. A method as claimed in claim 39 wherein the nucleic acid is expressed under the control of an inducible promoter.

41. A method for downwardy modulating the expression of a nucleic acid as claimed in any one of claims 1 to 12 in a plant, the method comprising any of the following:

(i) causing or allowing transcription from a nucleic acid as claimed in claim 13 in the plant;
(ii) causing or allowing transcription from a nucleic acid as claimed in claims 1 to 12 or a part thereof in the plant such as to reduce expression by co-suppression;
(iii) use of nucleic acid encoding a ribozyme specific for a nucleic acid as claimed in any one of claims 1 to 12.

42. A method of influencing any one or more of the following a phenotypic characteristics of a plant: insect protection; chemical production; climate tolerance; salt removal; fibre production; ornamental value; water and nutrient absorption; initiation of seed germination; pigmentation; taste; speed of seedling growth; the method comprising a method as claimed in any one of claims 39 to 41.

43. An isolated nucleic acid molecule comprising a sequence encoding the promoter sequence of the TGG1 gene, or a mutant, variant, derivative or other homolog thereof.

44. A vector comprising the promoter sequence of claim 43.

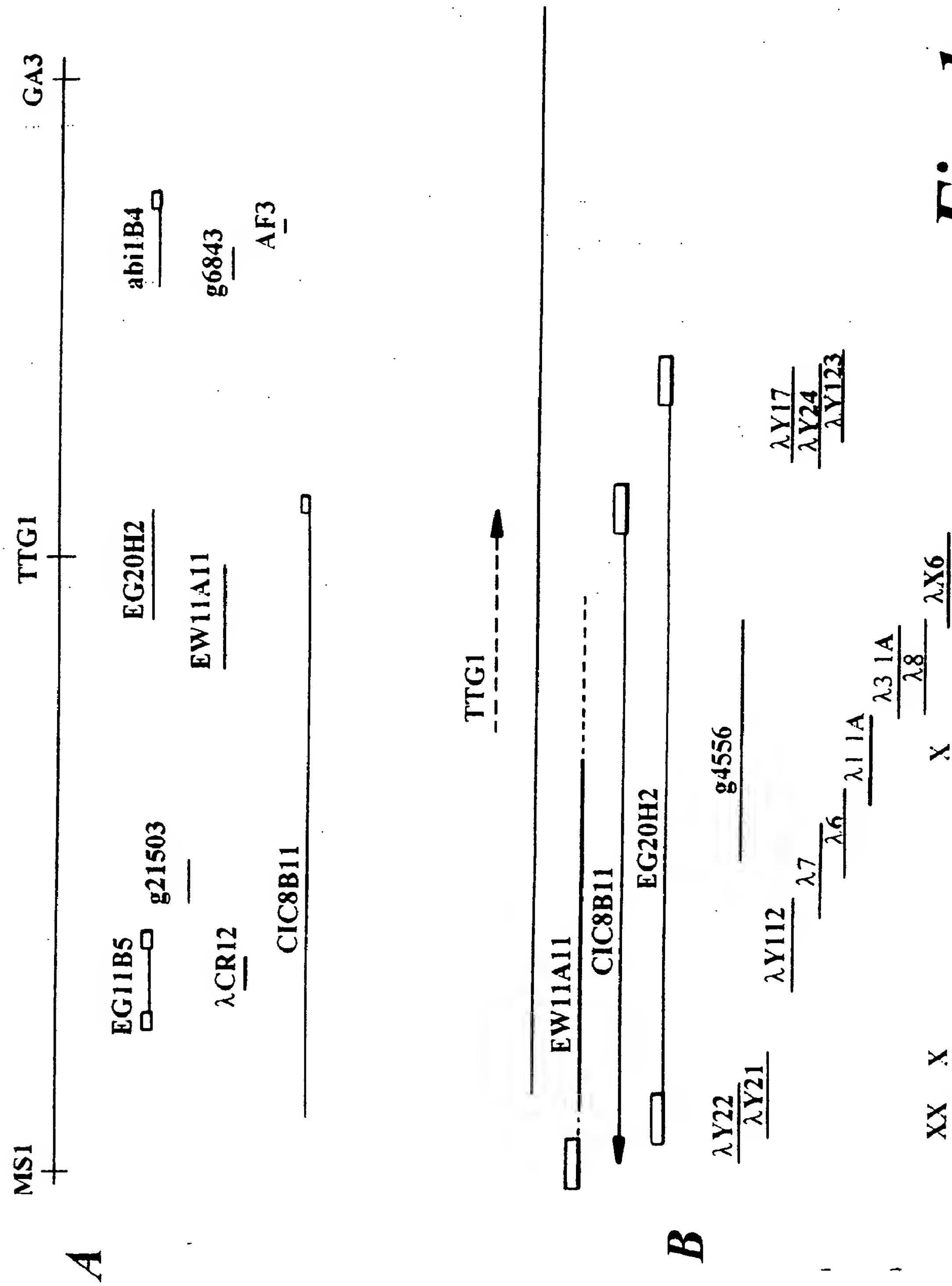
A
MS1
TTG1
GA3
EG11B5
g21503
λCR12
CIC8B11
EG20H2
EW11A11
abi1B4
g6843
AF3
B
TTG1
EW11A11
CIC8B11
EG20H2
g4556
λY22
λY21
λY112
λ7
λ6
λ1 1A
λ3 1A
λ8
λX6
λY17
λY24
λY123
XX
X
X
Crossover
In Dennis

Fig. 1

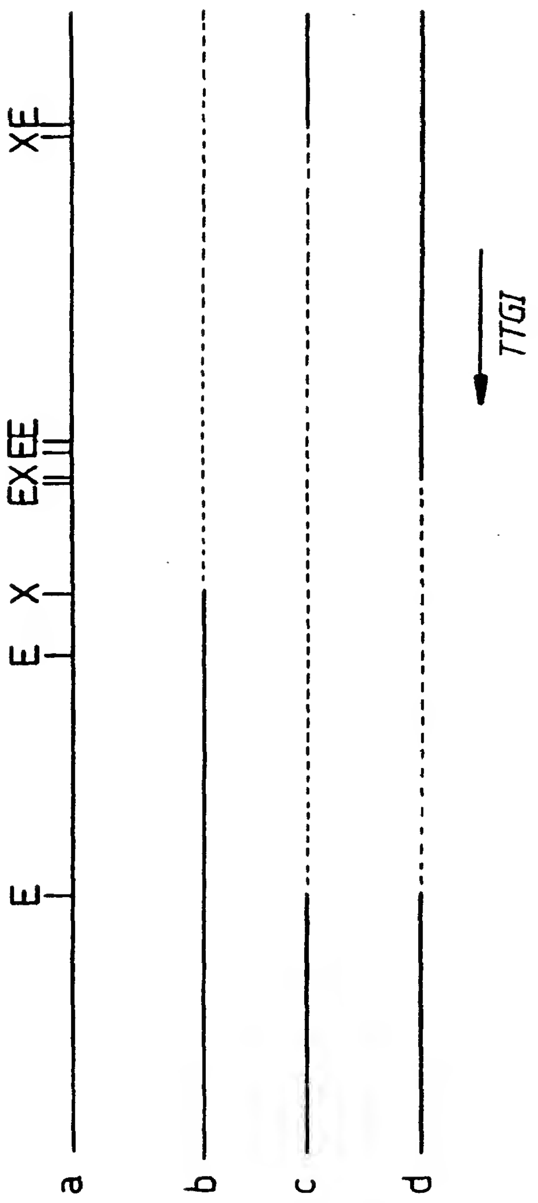

Fig. 2

Fig. 3A


```
agttatgtttatgcttttcatcatatttagtggttagtttttattatttatttattgatt    60
catgacttatgctagattatgataagaatttatgttaccacttgataaatcctccatttg   120
acatgtgtttaatgctagatttatattgtctccaaatttacaactttgatgtcttatgat   180
aaatgccaacaaccaaatttcagataaagattagcagactaactaagcttattattcact   240
tgcaaggtggagtgatgttgaaagaaccctcacagacacgtcattgggaagactaaatct   300
ctttttagcacgttacacctttgagatcgcgtttattccatatggagagagagcaacaat   360
acgagacatggagaggcaccattaccgccggcgcaactgcttccaaatattgacaaacaa   420
atttgaatctggatcttctctattcgtgaacaaggagatagaagctacgatgaatgcatg   480
gaagcttggtttgctttaatataaacactaaggggagtagaactttcttgaaaaattgt   540
cttcgaaccaaacgaaattatatttgtgatttcccctcatcttgaaagaacttttaaca   600
atgcaaattatttaccgaatgttaaaagctttttcgaataaattttacattttcttaat   660
aataataataaaaaggattgttgattatcttaatcacaaacaatttattttagctgaat   720
tagacaattgttagtaaaatgattagagtgtcacatattaatgttgttagtgtttcatgt   780
catcctagtgatccaataattaggccattctatagctcgtaacgttaaaataaaaggccc   840
attatctgaatatacagaagcccattatcaatagatacattaaaagatactgattaatcc   900
agagggtttatatctacgccgtctccattgattatttctccGTCTCTTGAAAAATCCGAC   960

                                           A in ttg1.10
TGACACTGACCTCAAAACTCTCCTCTCACTTTCGTCGTGAAGAAGCCAAATCTCGAATCG  1020
AATCAGCACCACACATTTCCATGGATAATTCAGCTCCAGATTCGTTATCCAGATCGGAAA  1080
                   M  D  N  S  A  P  D  S  L  S  R  S  E  T

                                                 G in ttg1.20
CCGCCGTCACATACGACTCACCATATCCACTCTACGCCATGGCTTTCTCTTCTCTCCGCT  1140
 A  V  T  Y  D  S  P  Y  P  L  Y  A  M  A  F  S  S  L  R  S

CATCCTCCGGTCACAGAATCGCCGTCGGAAGCTTCCTCGAAGATTACAACAACCGCATCG  1200
 S  S  G  H  R  I  A  V  G  S  F  L  E  D  Y  N  N  R  I  D

ACATTCTCTCTTTCGATTCCGATTCAATGACCGTTAAGCCTCTCCCGAATCTCTCCTTCG  1260
 I  L  S  F  D  S  D  S  M  T  V  K  P  L  P  N  L  S  F  E

AGCATCCTTATCCTCCAACAAAGCTAATGTTCAGTCCTCCTTCTCTCCGTCGTCCTTCCT  1320
 H  P  Y  P  P  T  K  L  M  F  S  P  P  S  L  R  R  P  S  S

CCGGAGATCTCCTCGCTTCCTCCGGCGATTTCCTCCGTCTTTGGGAAATTAACGAAGATT  1380
 G  D  L  L  A  S  S  G  D  F  L  R  L  W  E  I  N  E  D  S

CATCAACCGTCGAGCCAATCTCGGTTCTCAACAACAGCAAAACGAGCGAGTTTTGTGCGC  1440
 S  T  V  E  P  I  S  V  L  N  N  S  K  T  S  E  F  C  A  P

CGTTGACTTCCTTCGATTGGAACGATGTAGAGCCGAAACGTCTCGGAACTTGTAGTATTG  1500
 L  T  S  F  D  W  N  D  V  E  P  K  R  L  G  T  C  S  I  D

ATACGACGTGTACGATTTGGGATATTGAGAAGTCTGTTGTTGAGACTCAGCTTATAGCTC  1560
 T  T  C  T  I  W  D  I  E  K  S  V  V  E  T  Q  L  I  A  H

                          A in ttg1.19
ATGATAAAGAGGTTCATGACATTGCTTGGGGAGAAGCTAGGGTTTTCGCATCAGTCTCTG  1620
 D  K  E  V  H  D  I  A  W  G  E  A  R  V  F  A  S  V  S  A

CTGATGGATCCGTTAGGATCTTTGATTTACGTGATAAGGAACATTCTACAATCATTTACG  1680
 D  G  S  V  R  I  F  D  L  R  D  K  E  H  S  T  I  I  Y  E
```

4/7

Fig 3B

```
AGAGTCCTCAGCCTGATACGCCTTTGTTAAGACTTGCTTGGAACAAACAAGATCTTAGAT  1740
  S  P  Q  P  D  T  P  L  L  R  L  A  W  N  K  Q  D  L  R  Y

ATATGGCTACGATTTTGATGGATTCTAATAAGGTTGTGATTCTCGATATTCGTTCGCCGA  1800
  M  A  T  I  L  M  D  S  N  K  V  V  I  L  D  I  R  S  P  T

CTATGCCTGTTGCTGAGCTTGAAAgACATCAGGCTAGTGTGAATGCTATAGCTTGGGCGC  1860
  M  P  V  A  E  L  E  R  H  Q  A  S  V  N  A  I  A  W  A  P

                                    T in ttg1.9
CTCAGAGCTGTAAACATATTTGTTCTGGTGGTGATGATACACAGGCTCTTATTTGGGAGC  1920
  Q  S  C  K  H  I  C  S  G  G  D  D  T  Q  A  L  I  W  E  L

TTCCTACTGTTGCTGGACCCAATGGGATTGATCCGATGTCGGTTTATTCGGCTGGTTCGG  1980
  P  T  V  A  G  P  N  G  I  D  P  M  S  V  Y  S  A  G  S  E

        T in ttg1.21
AGATTAATCAGTTGCAGTGGTCTTCTTCGCAGCCTGATTGGATTGGTATTGCTTTTGCTA  2040
  I  N  Q  L  Q  W  S  S  S  Q  P  D  W  I  G  I  A  F  A  N

ACAAAATGCAGCTCCTTAGAGTTTGAGgtgagagtttctctttcgctacataattctcat  2100
  K  M  Q  L  L  R  V  *

ttgctaggcctagattctaatgaggaagcattgattattggtttagattgtgttgcatta  2160
cagatagttctctaggtttggtaactaaacgtttttcgattcttgataacaaagccact  2220
agagatttgacactaactcgttttagatttacctgaatcaatatctctgttaaaatcaat  2280
tactttgttatgcatacataaatcacagtttagtagtcatatatattggctcttattagc  2340
gacaggtctcacacttgctgtaatggctgatagtgtagtagtcatatgttggctttcatc  2400
taagttgatgtatcatatgatgaatagttgtacactcgtcaggttctaatttttacccat  2460
aattcttcagtctatttttttttgagacaatctattcttaatttaacgaagccactagct  2520
acgtatacaatattgttaatttaacgaagtatctgagaattgtttactgctgactctgc  2580
tgtatgccctcagaaacatatagaagtggaattggaaacttcatgctggtttgaacatct  2640
ttgtatgtgtgcttcaggtttttgtaactcatttagacaacagcattgcatatatacacg  2700
cacatatgcaacctagaaaatcaaataacctttccttataattactatccatttcacttg  2760
atgtcagGTGCAGATGTGAAGTGATCAATAAGGATTTTAGCATAGACCCGTATAATCGTC  2820
ATGTGCGTAAGTAGGTTTGGTTTGCGCTCCCTCTCGCTTTTAGGTCCGCAATGACTCTGT  2880
ATCTATCTGATTGTAACTAAAACTGAATTCATTTGATGAACCAAATGATACTATTATCTT  2940
ATGTTGTgtataaaacccaaccaggatatattgcggtttctggtgtttagatttggtaat  3000
tggagcttagtacaatgcaaccctgtcttgctttattggacgtctctaagataaatcagc  3060
```

5/7

```
                1                                                          50
        TTG1   -MDNSAPDSL SRSETAVTYD SPYPLYAMAF SSLRSSSGHR IAVGSFLEDY
   Matthiola   ---------- ---------- ---------- ---------- ----------
    Tobacco1   ---------- ---------- ---------- ---------- ----------
    Tobacco2   ---------- ---------- ---------- ---------- ----------
 Petuniaan11   MENSSQESQH LRSENSVTYD STYPIYSMAF SSF.PTPRRR IAVGSFIEEL

               51                                                        100
        TTG1   NNRIDILSFD SDSMTVKPLP NLSFEHPYPP TKLMFSPPSL RRPSSGDLLA
   Matthiola   ---------- ---------- ---FEHPYPP TKLMFSPPSL RRPSGGDLLA
    Tobacco1   ---------- ---------- ---FEHPYPP TKLMFHPNPS ASLKSNDILA
    Tobacco2   ---------- ---------- ---FEHPYPP TKLMFHPNPS ASLKSNDILA
 Petuniaan11   NNRVELLSFN EETLTLNPIP NLSFDHPYPP TKLMFHPNPI KS..NNDILA

               101                                                       150
        TTG1   SSGDFLRLWE INEDSSTVEP ISVLNNSKTS EFCAPLTSFD WNDVEPKRLG
   Matthiola   SSGDFLRLWE VSEDSSTVEP VSVLNNSKTS EFCAPLTSFD WNDVEPKRLG
    Tobacco1   SSGDYLRLWE VRE..SSIEP LFTLNNSKTS EYCAPLTSFD WNEVEPRRIG
    Tobacco2   SSGDYLRLWE VRE..SSIEP LFTLNNSKTS EYCAPLTSFD WDEIEPKRIG
 Petuniaan11   SSGDYLRLWE VKE..SSIEP LFTLNNSKTS EYCAPLTSFD WNEVEPKRIG

               151                                                       200
        TTG1   TCSIDTTCTI WDIEKSVVET QLIAHDKEVH DIAWGEARVF ASVSADGSVR
   Matthiola   TCSIDTTCTI WDIEKSVVET QLIAHDKEVH DIAWGEARVF ASVSADGSVR
    Tobacco1   TSSIDTTCTI WDVEKGVVQT QLIAHDKEGY DIAWGEAGVF ASVSADGSVR
    Tobacco2   TSSTDTTCTI WDVEKGVVET QLIAHDKEVY DIAWGEDGVF ASVSADGSVR
 Petuniaan11   TSSIDTTCTI WDVEKGVVET QLIAHDKEVY DIAWGEAGVF ASVSADGSVR

               201                                                       250
        TTG1   IFDLRDKEHS TIIYESPQPD TPLLRLAWNK QDLRYMATIL MDSNKVVILD
   Matthiola   IFDLRDKEHS TIIYESPQPD TPLLRLAWNK QDLRYMATIL MDSNKVVILD
    Tobacco1   IFDLRDKEHS TIIYESPQPD TPLLRVAWNK QDLRYMATIL MDSNKNVILD
    Tobacco2   IFDLRDKEHS TIIYESPQPD TPLLRLAWNK QDLRYMATIL MDSNKIVILD
 Petuniaan11   IFDLRDKEHS TIIYESPTPD TPLLRLAWNK QDLRYMATIL MDSNKVVILD

               251                                                       300
        TTG1   IRSPTMPVAE LERHQASVNA IAWAPQSCKH ICSGGDDTQA LIWELPTVAG
   Matthiola   IRSPTMPVAE LERHQASVNA IAWAPQSCKH ICSAGDDTQA LIWELPTVAG
    Tobacco1   IRSPAMPVAE LERHQASVNA IAWAPQSRRH ICSAGDDGQA LIWELPTV--
    Tobacco2   IRSPAMPVAE LERHQASVNA IAWAPQSCRH ICSAGDDGQA LIWELPTVAG
 Petuniaan11   IRSPAMPVAE LERHQASVNA IAWAPQSCRH ICSGGDDGQA LIWELPTVAG

               301                                        343
        TTG1   PNGIDPMSVY SAGSEINQLQ WSSSQPDWIG IAFANKMQLL RV-
   Matthiola   PNGIDP---- ---------- ---------- ---------- ---
    Tobacco1   ---------- ---------- ---------- ---------- ---
    Tobacco2   ---------- ---------- ---------- ---------- ---
 Petuniaan11   PNGIDPMSMY SAGAEINQLQ WSPAQRDWIA IAFSNKLQLL KV*
```

Fig 4

1	TTATTTCTCC	GTCTCTTGAA	AAATCCGACT	GACACTGACC	TCAAAACTCT
51	CCTCTCACTT	TCGTCGTGAA	GAAGCCAAAT	CTCGAATCGA	ATCAGCACCA
101	CACATTTCCA	TGGATAATTC	AGCTCCAGAT	TCGTTATCCA	GATCGGAAAC
151	CGCCGTCACA	TACGACTCAC	CATATCCACT	CTACGCCATG	GCTTTCTCTT
201	CTCTCCGCTC	ATCCTCCGGT	CACAGAATCG	CCGTCGGAAG	CTTCCTCGAA
251	GATTACAACA	ACCGCATCGA	CATTCTCTCT	TTCGATTCCG	ATTCAATGAC
301	CGTTAAGCCT	CTCCCGAATC	TCTCCTTCGA	GCATCCTTAT	CCTCCAACAA
351	AGCTAATGTT	CAGTCCTCCT	TCTCTCCGTC	GTCCTTCCTC	CGGAGATCTC
401	CTCGCTTCCT	CCGGCGATTT	CCTCCGTCTT	TGGGAAATTA	ACGAAGATTC
451	ATCAACCGTC	GAGCCAATCT	CGGTTCTCAA	CAACAGCAAA	ACGAGCGAGT
501	TTTGTGCGCC	GTTGACTTCC	TTCGATTGGA	ACGATGTAGA	GCCGAAACGT
551	CTCGGAACTT	GTAGTATTGA	TACGACGTGT	ACGATTTGGG	ATATTGAGAA
601	GTCTGTTGTT	GAGACTCAGC	TTATAGCTCA	TGATAAAGAG	GTTCATGACA
651	TTGCTTGGGG	AGAAGCTAGG	GTTTTCGCAT	CAGTCTCTGC	TGATGGATCC
701	GTTAGGATCT	TTGATTTACG	TGATAAGGAA	CATTCTACAA	TCATTTACGA
751	GAGTCCTCAG	CCTGATACGC	CTTTGTTAAG	ACTTGCTTGG	AACAAACAAG
801	ATCTTAGATA	TATGGCTACG	ATTTTGATGG	ATTCTAATAA	GGTTGTGATT
851	CTCGATATTC	GTTCGCCGAC	TATGCCTGTT	GCTGAGCTTG	AAAgACATCA
901	GGCTAGTGTG	AATGCTATAG	CTTGGGCGCC	TCAGAGCTGT	AAACATATTT
951	GTTCTGGTGG	TGATGATACA	CAGGCTCTTA	TTTGGGAGCT	TCCTACTGTT

Fig. 5A

```
1001  GCTGGACCCA ATGGGATTGA TCCGATGTCG GTTTATTCGG CTGGTTCGGA
1051  GATTAATCAG TTGCAGTGGT CTTCTTCGCA GCCTGATTGG ATTGGTATTG
1101  CTTTTGCTAA CAAAATGCAG CTCCTTAGAG TTTGAGGTGC AGATGTGAAG
1151  TGATCAATAA GGATTTTAGC ATAGACCCGT ATAATCGTCA TGTGCGTAAG
1201  TAGGTTTGGT TTGCGCTCCC TCTCGCTTTT AGGTCCGCAA TGACTCTGTA
1251  TCTATCTGAT TGTAACTAAA ACTGAATTCA TTTGATGAAC CAAATGATAC
1301  TATTATCTTA TGTTGTGTAT AAAACCCAAC CAGGATATAT TGCGGTTTCT
1351  GGTGTTTAGA TTTGGTAATT GGAGCTTAGT ACAATGCAAC CCTGTCTTGC
1401  TTTATTGGAC GTCTCTAAGA TAAATCAGC
```

Fig. 5B

INTERNATIONAL SEARCH REPORT

In[illegible]tional Application No
PCT/GB 98/01861

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/29 C12N15/82 C12N15/11 C12N5/10 C12Q1/68
C07K14/415 C07K16/16 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N C12Q A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	COOKE R. ET AL.: "AC F20055" EMBL DATABASE, 5 March 1996, XP002082374 HEIDELBERG see the whole document ---	13
X	COOKE R. ET AL.: "AC F20056" EMBL DATABASE, 5 March 1996, XP002082375 Heidelberg cited in the application see the whole document ---	4,5,14, 16,19, 23-25

-/--

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 October 1998	11/11/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Kania, T

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	LLOYD A. ET AL.: "Arabidopsis and Nicotiana anthocyanin production activated by maize regulators R and C1" SCIENCE, vol. 258, 11 December 1992, pages 1773-1775, XP002003019 cited in the application * see esp. p.1774/75 * ---	1,11,13, 19-32, 34-42
X	DE VETTEN N. ET AL.: "The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals" GENES & DEVELOPMENT, vol. 11, no. 11, 1 June 1997, pages 1422-1434, XP002082376 cited in the application see the whole document ---	4,10, 19-32, 34-42
A	SHIRLEY B. ET AL.: "Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis" THE PLANT JOURNAL, vol. 8, no. 5, November 1995, pages 659-671, XP002082377 cited in the application * see the whole document, esp. discussion, figure 2 * ---	1-44
A	LARKIN J C ET AL: "Roles of the GLABROUS1 and TRANSPARENT TESTA GLABRA genes in Arabidopsis trichome development." PLANT CELL 6 (8). 1994. 1065-1076. ISSN: 1040-4651, XP002004687 cited in the application see the whole document ---	1-44
A	GALWAY M. ET AL.: "The TTG gene is required to specify epidermal cell fate and cell patterning in the Arabidopsis root" DEVELOPMENTAL BIOLOGY, vol. 166, no. 2, December 1994, pages 740-754, XP002082378 cited in the application see the whole document -----	1-44

1